

# Introduction

## Artificial Intelligence, Machine Learning and Deep Learning

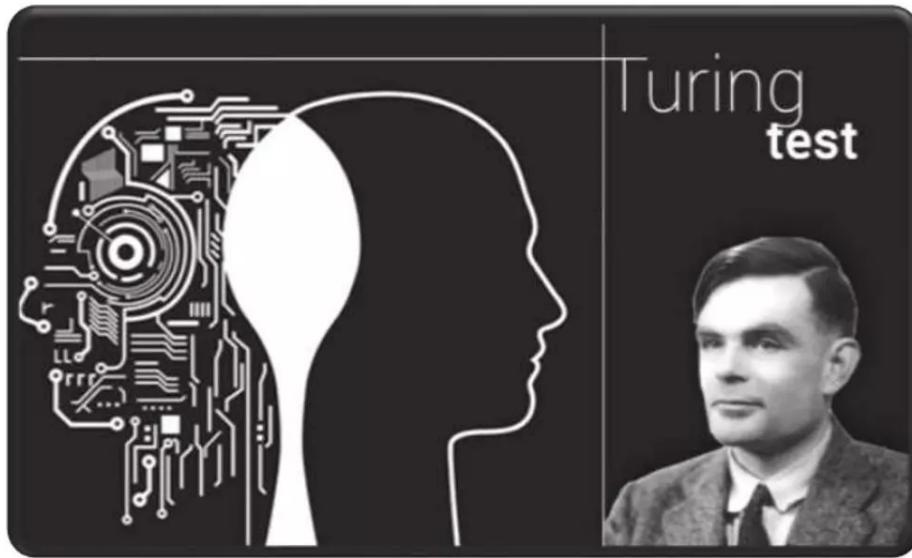
- The term **AI** in the broadest sense refers to simulation of human intelligence processes by computer systems
- **Machine Learning** is a subset of AI focusses on designing specific systems which can learn and make decisions/predictions based on data.
- **Deep Learning** is a subset of Machine Learning that uses a specific set of algorithms known as Neural-Networks often with many layers.



# WHAT IS A.I.?

- 1) The study of computer systems that attempt to model and apply the intelligence of the human mind.
- 2) A branch of computer science dealing with the simulation of intelligent behaviour in computers.
- 3) The capability of a machine to imitate intelligent human behaviour.

# Early History



- In 1950 English mathematician Alan Turing wrote a landmark paper titled “Computing Machinery and Intelligence” that asked the question: “Can machines think?”
- Further work came out of a 1956 workshop at Dartmouth sponsored by John McCarthy. In the proposal for that workshop, he coined the phrase a “study of Artificial Intelligence”

# CURRENT STATUS OF AI

- ❖ AI have taken many shapes and forms over recent years
  - Mobile Phones ( Siri/ Cortana)
  - Video Games Characters
  - GPS/ Voice Recognition
  - Robotics
  
- ❖ Google has been a major play on AI transcendence and Deep Learning.
  - Deep learning is a machine learning based on algorithms.

# Introduction

## Artificial Intelligence, Machine Learning and Deep Learning

- The term **AI** in the broadest sense refers to simulation of human intelligence processes by computer systems
- **Machine Learning** is a subset of AI focusses on designing specific systems which can learn and make decisions/predictions based on data.
- **Deep Learning** is a subset of Machine Learning that uses a specific set of algorithms known as Neural-Networks often with many layers.

## Traditional Programming



## Machine Learning



Supervised

Unsupervised

Types of  
Machine  
Learning

Semi-supervised

Reinforcement

# Model

- A model is the relationship between features and the label.
- An ML model is a mathematical model that generates predictions by finding patterns in your data.
- ML Models generate predictions using the patterns extracted from the input data.
- A model represents what was learned by a machine learning algorithm.
- The model is the “thing” that is saved after running a machine learning algorithm on training data and represents the rules, numbers, and any other algorithm-specific data structures required to make predictions.

# Introduction

## Types of Machine Learning Models

### □ Supervised Learning

Supervised Learning is a type of Machine Learning model trained on labeled data

### Email Spam Classification Model

Data: Examples of emails either tagged as Spam or not Spam

#### Training:

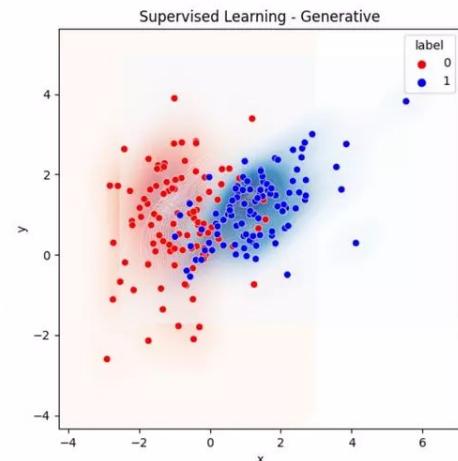
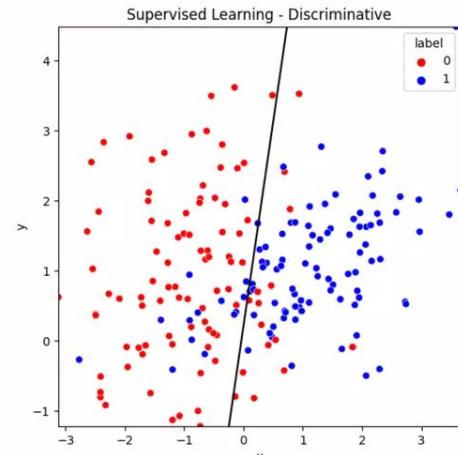
**Discriminative** – Learns the boundary that separates “spam” vs “not spam”

**Generative** – Learns the distribution of “spam” and “not spam” emails to understand how each class generates content

#### Inference

**Discriminative** – Determine on which side of the boundary a new email falls

**Generative** – Based on learned distributions compute the likelihood of the new email being “spam” vs “not spam”



# Introduction

## Types of Machine Learning Models

### Unsupervised Learning

Unsupervised Learning is a type of Machine Learning model that identifies patterns and structures within un-labelled data

### Email Topic Modeling

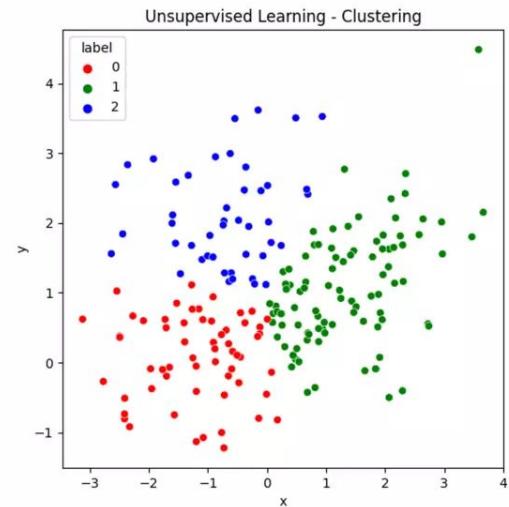
Data: A large collection of emails you may want to organize by subject matter

### Training:

Learn the distribution that generates the structure within the data

### Inference

- Assign new email to the cluster where they have the highest probability of belonging



# Introduction

## Types of Machine Learning Models

### Reinforcement Learning

Interaction: Agent interacts with the environment by choosing actions from its current policy

*A self-driving car decides to take a left or a right based on its current strategy and current state of the road*

Reward/Penalty: After each action, agent receives a reward/penalty which reflects the success of the action

*If the car safely navigates traffic or obeys rules, it's a success*

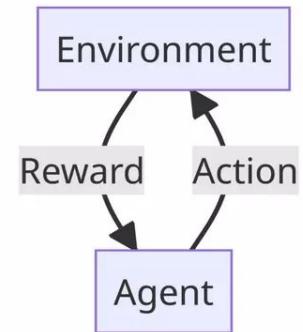
Policy Update: Agent updates the policy based on feedback received aiming to maximize the total reward over time.

*Based on the reward/penalty received car adjusts its driving policy, actions with positive rewards will be repeated and negative rewards will be avoided*

### Shallow and Deep Models

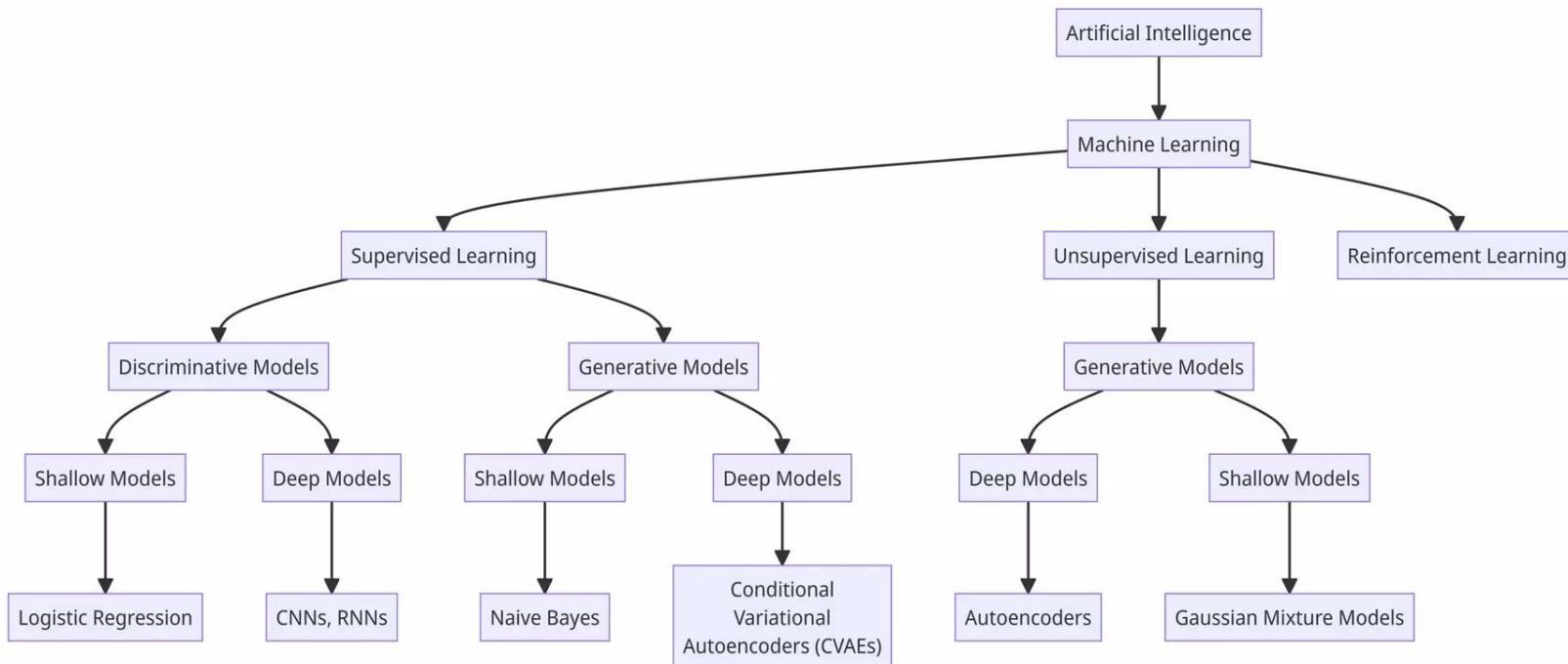
Models with limited layers and capable of capturing only linear and simple nonlinear relationships are called shallow models

Models with many layers and capable of capturing complex hierarchical patterns are called deep models



# Introduction

## Summary



## **Generative AI**

---

### **GPT, GAN and Diffusion Models**

---

### **Applications of Generative AI**

---

### **Emerging Trends, Limitations, Potential Ahead**

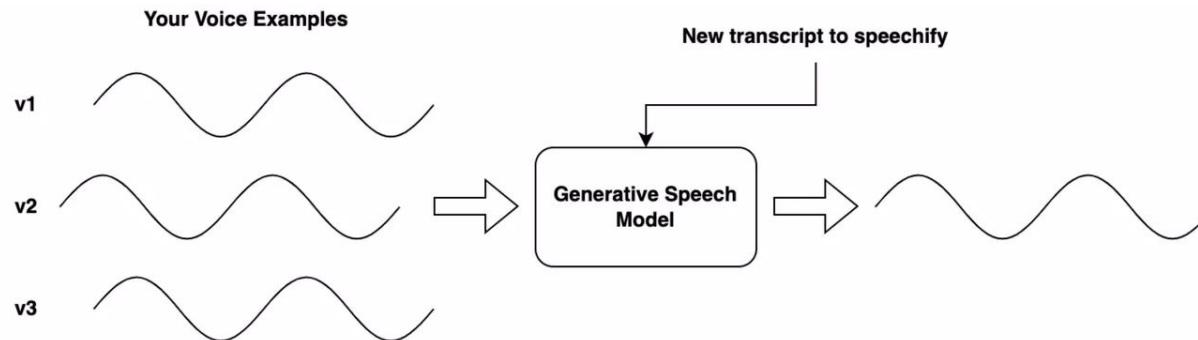
---

# Generative AI

## Definition

Generative AI refers to a set of artificial intelligence methodologies that can produce novel content that resembles the training data they were exposed to.

The content could be anything spanning from synthesizing text, generating code, realistic images, music and more



# Generative AI

## Generative Pre-Trained Transformer (GPT) - Motivation

### Issues with CNNs, RNNs, LSTMs

- Convolutional Neural Networks (**CNNs**) are good at **local feature extraction** and struggle to understand long-range dependencies in data
- CNNs **do not have a mechanism to understand the order of elements** making it harder for problems involving text and time-series
- **RNNs** especially **LSTMs** can handle long range dependencies due to their ability to process data sequentially. But as the sequences get longer, they struggle from **vanishing gradient problems**
- CNNs, RNNs, LSTMs are suitable for specific data types and are **not efficient at handling multi-modal inputs**

What if you can completely avoid recurrent connections, thereby avoiding vanishing gradient issues?

# Generative AI

## Generative Pre-Trained Transformer (GPT) - Motivation

- A new architecture called **Transformers** is proposed by scientists from Google which avoids the recurrent connections altogether by relying on an operation known as attention
- This architecture also takes care of sequential nature of inputs by using **positional embeddings**

### Attention Is All You Need

Ashish Vaswani\*  
Google Brain  
avaswani@google.com

Noam Shazeer\*  
Google Brain  
noam@google.com

Niki Parmar\*  
Google Research  
nikip@google.com

Jakob Uszkoreit\*  
Google Research  
usz@google.com

Llion Jones\*  
Google Research  
llion@google.com

Aidan N. Gomez\* †  
University of Toronto  
aidan@cs.toronto.edu

Lukasz Kaiser\*  
Google Brain  
lukaszkaiser@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

[https://proceedings.neurips.cc/paper\\_files/paper/2017/file/3f5ee243547dee91fdb053c1c4a845aa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fdb053c1c4a845aa-Paper.pdf)

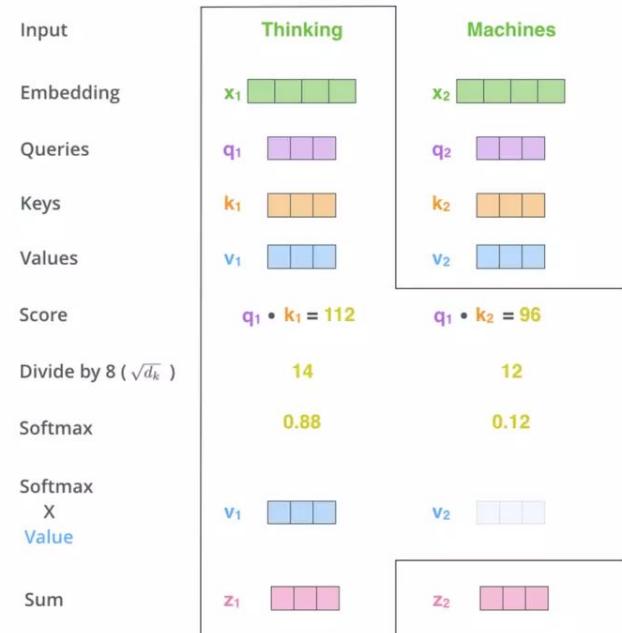
# Generative AI

## Generative Pre-Trained Transformer (GPT) - Attention

Let's take an example sentence.

Alice, who has a black cat, loves going to park

- When the model is processing the word "loves", attention mechanism allows it to associate it with "Alice"
  - At each word, attention mechanism allows to look at words at other positions in the input sequence to better encode the word at current position
1. At each input position, calculate query, key and value vectors (a linear transformation of embeddings using learnt weight matrices)
  2. Compute dot product between each query and all the keys in the input sequence (attention)
  3. Compute a weighted sum of all value vectors using attention weights as coefficients



<http://jalammar.github.io/illustrated-transformer/>

# Generative AI

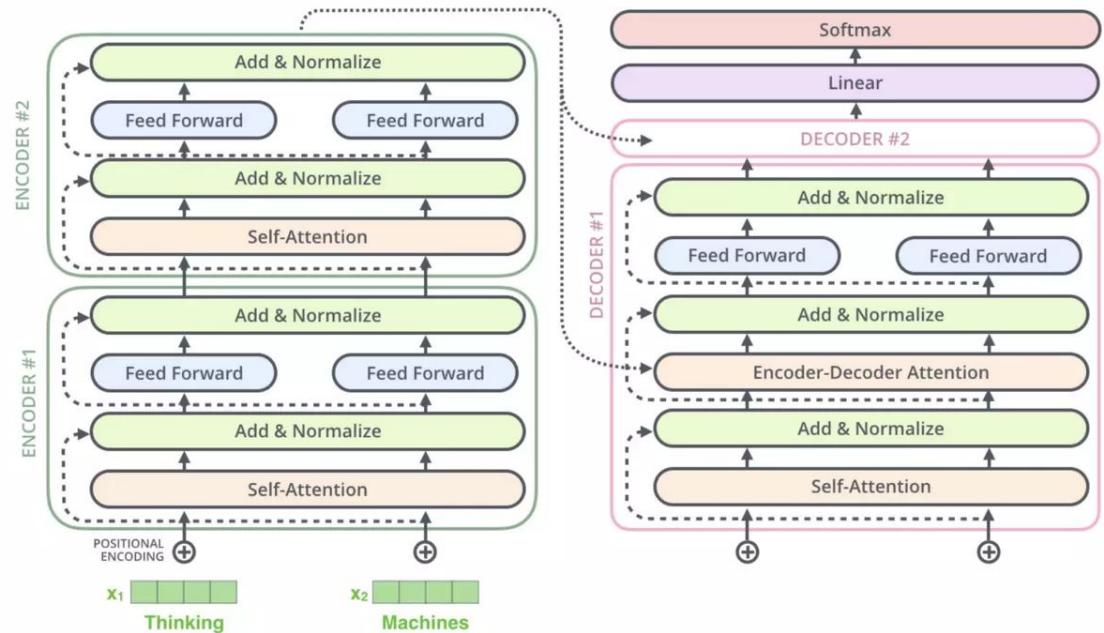
## Generative Pre-Trained Transformer (GPT) –Transformers Architecture

### Architecture

- Six Encoder layers stacked
- Six Decoder layers stacked
- Positional Embeddings
- Masked Attention (Encoder-Decoder Attention)

### Advantages

- Better long-range connections
- Easier to parallelize
- Can make the networks much deeper (more layers) than RNNs



<https://arxiv.org/pdf/1706.03762.pdf>

<http://jalammar.github.io/illustrated-transformer/>

<https://cs182sp21.github.io/static/slides/lec-12.pdf>

For Wifi use Greenline

# Generative AI

## Generative Pre-Trained Transformer (GPT)

A **Generative-Pre-Trained Transformer** is a kind of transformers model developed by OpenAI for natural language processing tasks

- **Generative** refers to the model's ability to generate text
- **Pre-Trained** refers to models training process consisting of two stages
  - **Pre-Training**: Model is trained on a large corpus of text data, where the objective is to predict next word in a sentence
  - **Fine-tuning**: Once the model is pre-trained the model can be fine-tuned on a specific task with a task-specific dataset with supervised learning

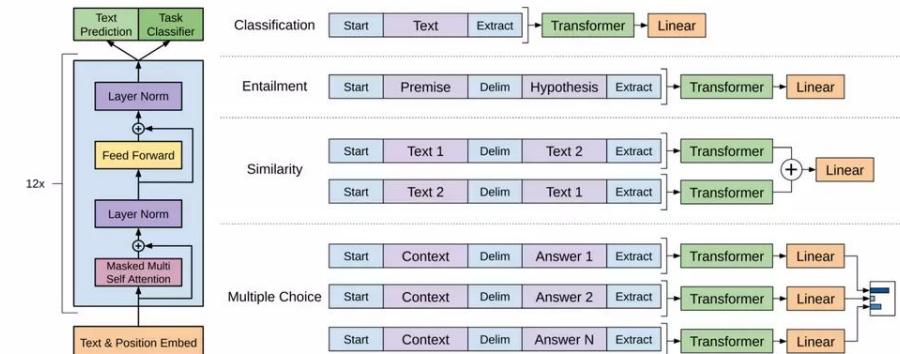


Figure 1: (left) Transformer architecture and training objectives used in this work. (right) Input transformations for fine-tuning on different tasks. We convert all structured inputs into token sequences to be processed by our pre-trained model, followed by a linear+softmax layer.

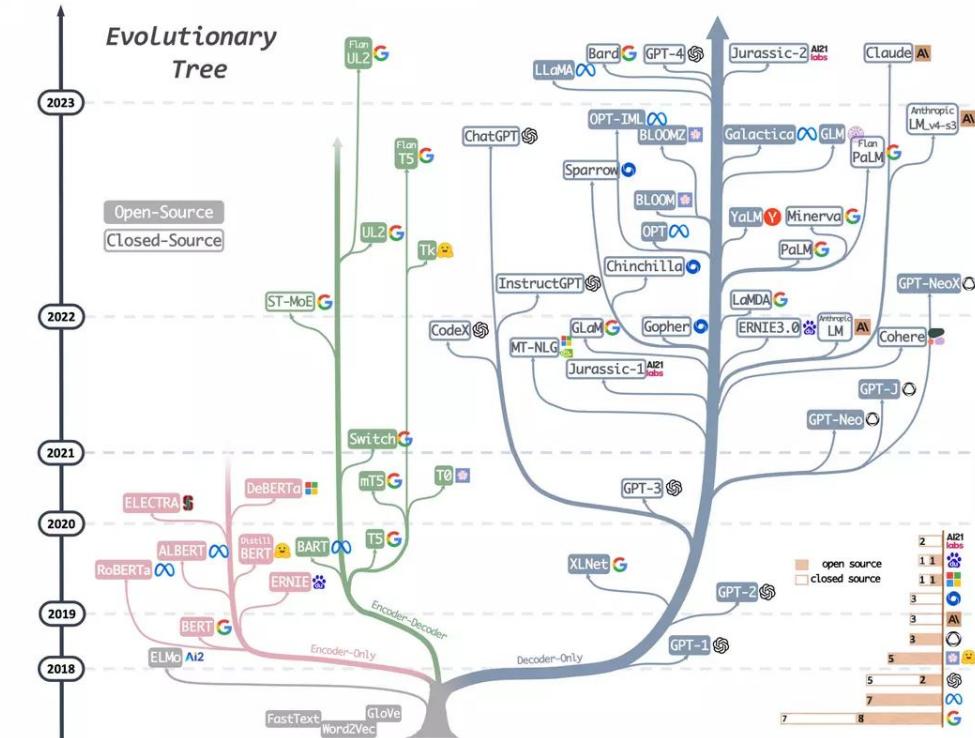
# Generative AI

## Generative Pre-Trained Transformer (GPT) – LLM Landscape

**Encoder Models:** These models map input sequences to a vector representation. Useful for extracting features (**BERT**)

**Decoder Models:** These models generate an output sequence from a fixed length input vector. Useful for generation text, images etc. (**GPT-3**)

**Encoder-Decoder Models:** These models are a combination of both encoder and decoder. Encoder is responsible for mapping input into vector and decoder generates output sequence from that vector. (**BART/ T5/ FLAN UL2**)



<https://arxiv.org/pdf/2304.13712.pdf>

<https://amatriain.net/blog/transformer-models-an-introduction-and-catalog-2d1e9039f376/>

For Wifi use Greenline

# Generative AI

## Generative Pre-Trained Transformer (GPT) – Chain-of-thought Prompting

**Chain-of-Thought Prompting** is a technique that enables LLMs to complex reasoning generating a chain-of-thought, a series of intermediate reasoning steps.

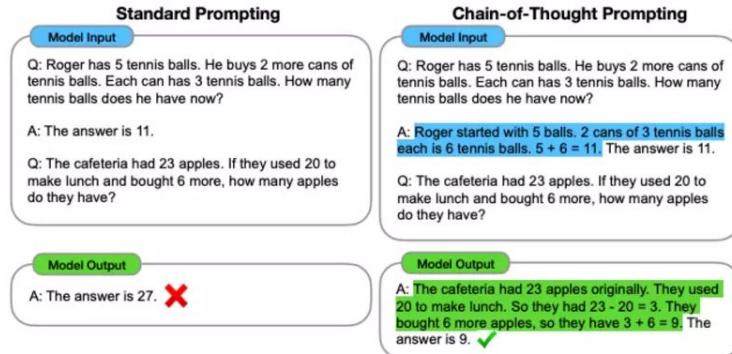


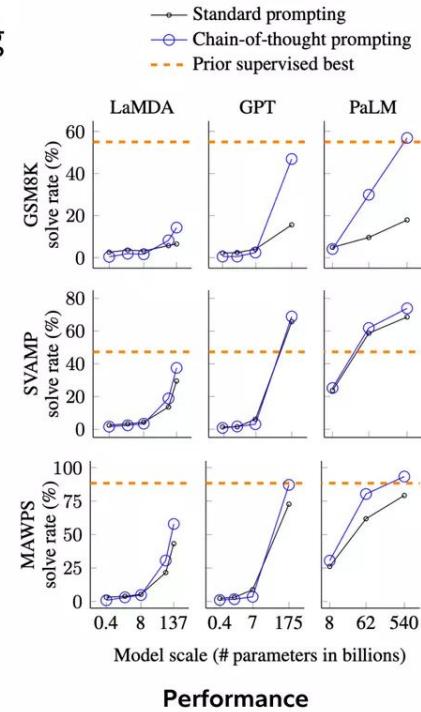
Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

### Prompting

Table 12: Summary of math word problem benchmarks we use in this paper with examples.  $N$ : number of evaluation examples.

Dataset	$N$	Example problem
GSM8K	1,319	Josh decides to try flipping a house. He buys a house for \$80,000 and then puts in \$50,000 in repairs. This increased the value of the house by 150%. How much profit did he make?
SVAMP	1,000	Each pack of dvds costs \$76 dollars. If there is a discount of 25 dollars on each pack. How much do you have to pay to buy each pack?
ASDiv	2,096	Ellen has six more balls than Marin. Marin has nine balls. How many balls does Ellen have?
AQuA	254	A car is being driven, in a straight line and at a uniform speed, towards the base of a vertical tower. The top of the tower is observed from the car and, in the process, it takes 10 minutes for the angle of elevation to change from $45^\circ$ to $60^\circ$ . After how much more time will this car reach the base of the tower? Answer Choices: (a) $5\sqrt{3} + 1$ (b) $6\sqrt{3} + \sqrt{2}$ (c) $7\sqrt{3} - 1$ (d) $8\sqrt{3} - 2$ (e) None of these
MAWPS: SingleOp	562	If there are 7 bottle caps in a box and Linda puts 7 more bottle caps inside, how many bottle caps are in the box?
MAWPS: SingleEq	508	Benny bought a soft drink for 2 dollars and 5 candy bars. He spent a total of 27 dollars. How much did each candy bar cost?
MAWPS: AddSub	395	There were 6 roses in the vase. Mary cut some roses from her flower garden. There are now 16 roses in the vase. How many roses did she cut?
MAWPS: MultiArith	600	The school cafeteria ordered 42 red apples and 7 green apples for students lunches. But, if only 9 students wanted fruit, how many extra did the cafeteria end up with?

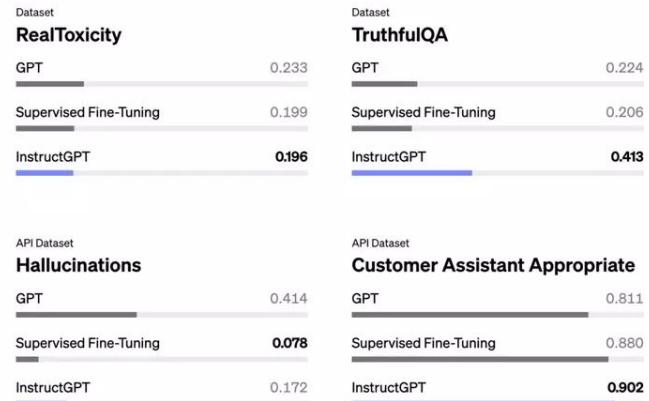
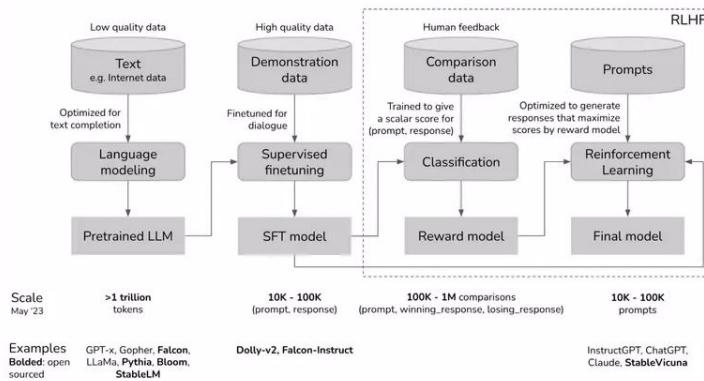
### Datasets and Example Problems



# Generative AI

## Generative Pre-Trained Transformer (GPT) – Alignment - RLHF

- **Reinforcement Learning through Human Feedback** is technique that allows models to learn directly from human feedback (like prompting) without the need for labeled data
- Due to the nature of training data being scrapped from internet (contains a lot of mis-information, conspiracy theories etc..) the models must be **further polished/aligned using RLHF** to make it user appropriate



Evaluating InstructGPT for toxicity, truthfulness, and appropriateness. Lower scores are better for toxicity and hallucinations, and higher scores are better for TruthfulQA and appropriateness. Hallucinations and appropriateness are measured on our API prompt distribution. Results are combined across model sizes.

# Generative AI

## Generative Adversarial Networks (**GANs**)

Imagine you have a bunch of **cat images**, and you want a machine learning model to create similar images. This is exactly what a GAN does.

**Generator:** Takes in random numbers as input and generates the images of interest (**the forger**)

**Discriminator:** Takes both the images from the generator and the real images from the data and spots the difference between them (**the detective**)

Both the generator and the discriminator are trained together. And, over the duration of training, the generator gets better at creating images which look real, and the discriminator gets better at spotting fakes.

**Adversarial Objective:** These two networks are pitted against each other where the **generator creates more realistic synthetic images to fool the discriminator** while the **discriminator networks tries to get better at detecting fake images**. This back-and-forth strategy forces both the networks to improve until the generator can create highly realistic synthetic images, that indistinguishable from real images

# Generative AI

## Diffusion Models

**Diffusion models** are another class of Generative models which work by adding noise to the images in the training data by a process called **forward diffusion process** and then reversing the process to recover the original image using **reverse diffusion**. These models can be trained on **large unlabeled datasets in an unsupervised manner**.

**Stable Diffusion:** Stable Diffusion is a **text-to-image model**. A stable diffusion model has four important elements

- **Diffusion Probabilistic Model**
- **U-Net Architecture**
- **Latent Text Encoding**
- **Classifier-Free Guidance**

# Applications

## Dall-E Example 2

S DALL-E History Collections

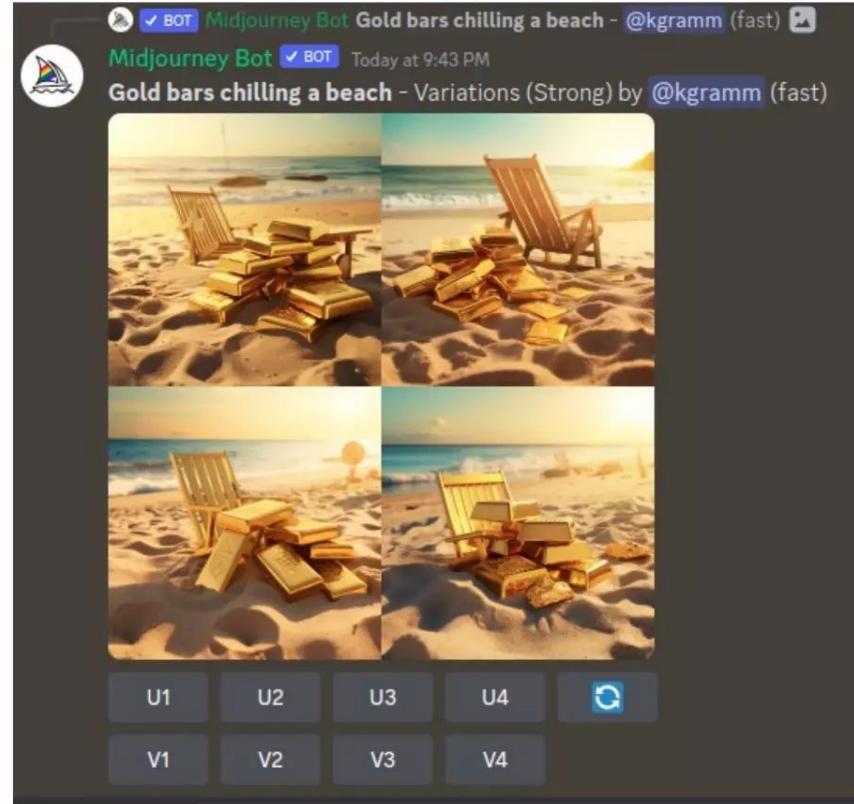
Edit the detailed description Surprise me Upload →

Gold bars chilling a beach Generate

The image displays four generated photographs of gold bars on a beach, arranged horizontally. Each photograph shows a different arrangement of gold bars on sandy ground. The first image shows several bars scattered. The second image shows three bars standing upright. The third image shows a large, single gold bar. The fourth image shows two bars resting on a rocky surface. The background in all images includes a beach, ocean waves, and a clear sky.

# Applications

## Midjourney

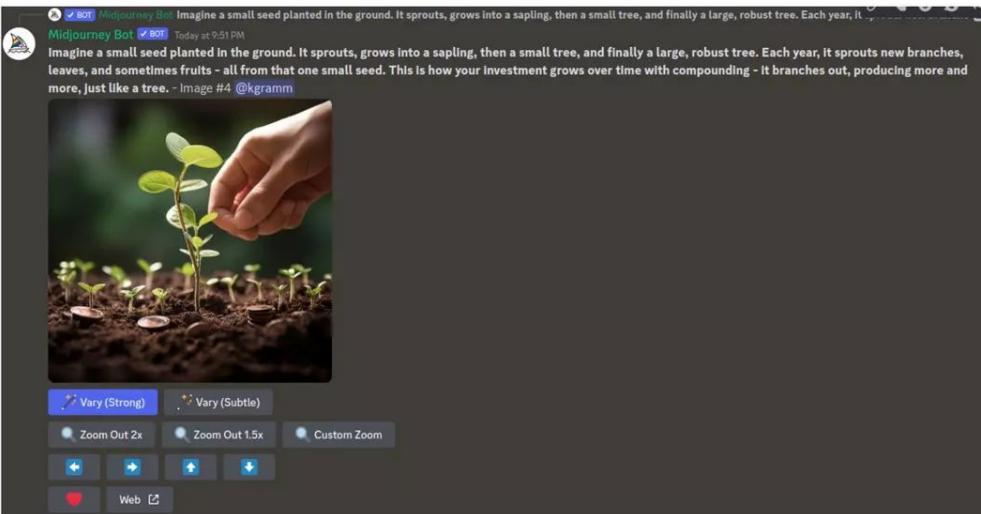


For Wifi use Greenline

# Applications

## Midjourney

**Prompt:** Imagine a small seed planted in the ground. It sprouts, grows into a sapling, then a small tree, and finally a large robust tree. Each year, it sprouts new branches, leaves and sometimes fruits – all from that small seed. This is how your investment grows with compounding – It branches out producing more and more just like a tree



The image shows the Midjourney application interface. At the top, there are two messages: one from 'Midjourney Bot' and one from 'Midjourney Bot'. Below the messages is a generated image of a small plant growing from a stack of coins in the soil. The plant has several green leaves and is reaching towards a hand that is gently touching its stem. The interface includes various control buttons at the bottom, such as 'Vary (Strong)', 'Zoom Out 2x', 'Custom Zoom', and 'Web'.



For Wifi use Greenline

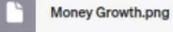
Images Generated for this presentation

# Applications

## ChatGPT

Model: Code Interpreter

KA Given this image can you create a nice gif out of it by adding a motion effect?

 Money Growth.png

KA Sure, I can help with that. Could you please specify the type of motion effect you'd like to apply to the image? For example, are you interested in a zoom in/out effect, a sliding effect, a rotation effect, or something else?

KA Zoom in effect please

Finished working Show work

I've created a GIF with a zoom-in effect. You can download it using the link below:

[Download the GIF](#)



For Wifi use Greenline

Images Generated for this presentation

# Generative AI

## Emerging Trends

- Major breakthroughs in deep learning architectures like Transformers and Generative Adversarial Networks
- Availability of massive datasets and GPU/TPU compute
- New advances in techniques like RLHF/Prompting made it much easy to align these models
- Low barrier of entry due to intuitive and user-friendly interfaces and strong open-source ecosystem
- GenAI holds potential to create photo-realistic images, human-like speech and text and generate working code from natural language descriptions which was not possible until recently

# Things to Keep in Mind

1. **Lack of Consistency (Hallucination)**: LLMs tend to produce wildly different answers, when the same question is asked multiple times
2. **Bias**: As the models are trained on data scrapped from internet, they might have inherited the biases present in the training data
3. **Interpretability**: It is difficult to understand why a particular response or content is generated, making it very challenging for use cases where explainability is inherently required.
4. **Real-time Knowledge**: As the models are trained on a fixed dataset at a particular point in time, they lack information/changes that occurred after that point.
5. **Memory**: Even though these models are getting good with context lengths that can be supported, having an efficient memory remembering the important details of conversations over a long period of time is still a challenging task.
6. **Engineering Challenges**: Operating these semi-non-deterministic models especially in a multi-model setting (including voice, text, images etc..) at scale remains a significant challenge

# Potential of Generative AI

1. **Low Resource Languages** – Ability to understand, generate any language, especially low resource ones, could help study languages and historical documents in general
2. **Inclusion and Accessibility** – Avatars proficient in sign languages, high precision caption generation etc., could increase accessibility for all people
3. **Personalized Content Generation** – Video games, music, movies can be created that cater to users and individual interests at scale
4. **AI Tutors** – Imagine a world where you can conjure up a tutor to teach you any skill you would like to learn at your own pace
5. **Intelligent Assistants** – Laborious and repetitive tasks can be delegated to Intelligent Assistants allowing humans to focus on critical thinking and decision making
6. **Accelerating Scientific Discovery**- General advances in AI can help accelerate scientific discovery by generating deep insights from massive datasets and design new algorithms. This can help solve most challenging problems we face today.

<https://www.forbes.com/sites/bernardmarr/2023/05/31/the-future-of-generative-ai-beyond-chatgpt/?sh=161c85da3da9>